The Cyberinfrastructure for Scholars Project: Componentized Architecture for Sustainable Scholarly Portals

By: Aaron Krowne, Stacey Martin, Urvashi Gadi, Micah Wedemeyer, Martin Halbert

"The Cyberinfrastructure for Scholars Project: Componentized Architecture for Sustainable Scholarly Portals," (with Aaron Krowne, Stacy Martin, Urvashi Gadi, and Micah Widemeyer) in *JCDL 2007: Proceedings of the Seventh ACM/IEEE Joint Conference on Digital Libraries: Building & Sustaining the Digital Environment*, Vancouver, June 18-23, 2007. Edited by Edie Rasmussen, Ray R. Larson, Elaine Toms, and Shigeo Sugimoto. ISBN: 9781595936448. p. 487.

Made available courtesy of ACM: http://dx.doi.org/10.1145/1255175.1255286

Abstract:

With the advent in recent years of widespread digital collections, web services (including "Web 2.0"), and metadata harvesting (most notably with OAI-PMH), a challenge has arisen of meshing these elements into coherent and sustaining scholarly information environments ("digital libraries," "portals," etc.). Many system instances have been successfully constructed demonstrating technical solutions to various problems in these areas, but these instances have not as of yet translated into wide adoption.

Keywords: Componentized Architecture | OXF | Web Services | OAI | Digital Libraries | Design | Management | Standardization

Article:

1. Introduction

With the advent in recent years of widespread digital collections, web services (including "Web 2.0"), and metadata harvesting (most notably with OAI-PMH), a challenge has arisen of meshing these elements into coherent and sustaining scholarly information environments ("digital libraries," "portals," etc.). Many system instances have been successfully constructed demonstrating technical solutions to various problems in these areas, but these instances have not as of yet translated into wide adoption.

For the Cyberinfrastructure for Scholars project (funded by the Mellon Foundation), Emory University is building the "SouthComb" Southern Studies scholarly portal as a capstone of a series of prior and ongoing grant projects (including MetaCombine, Quality Metrics, and American South, and MetaArchive) with the goal of extending the findings and advances of these projects into a real-world, sustaining scholarly portal.

Our approach with SouthComb is to lower the technical cost of maintenance through componentized architecture, while simultaneously engaging with the professional community to establish a companion organizational and societal support structure with an appropriate business model. With a lower maintenance cost, the pressures for revenue streams can be significantly reduced, mitigating many onerous intellectual property "ownership" dilemmas.

In this poster, we primarily address the technical side of our solution to the sustainability problem.

The high-level outline of this architecture is, briefly, as follows. Three key "tiers" of the system are componentized separately: (1) the collection development layer, (2) the user-facing service layer, and (3) the management layer. The first item can be thought of as "data production" componentization, and the latter two can be thought of as "componentized services." (See Figure 1 for a diagram).

Data production componentization is based on the OAI/OCKHAM transformations Framework (OXF) paradigm we are developing [2, 1, 3]. In this paradigm, collection production needs are broken down into discrete steps of OAI repository transformations, e.g., for tasks like classification, metadata normalization, geocoding, and more. A "pipeline" of such steps is arranged, at the end of which an enhanced metadata representation of the collection is produced, suitable for building services with considerably more depth than would otherwise be possible. All of these steps, by way of their companion OXF services, can be shared with others directly (over the internet as Web Services) or re-used as local software components.

Service componentization (both management and userfacing) is provided by the web services framework and methodology, invoking a Service Oriented Architecture (SOA). Each distinct digital library service will be implemented as a module in this scheme, drawing on records from a central (Fedora) repository, where the finalized collection will be kept. As "portlets," these services are more easily reusable in other information portals which can support web services, which will help foster uptake into a variety of systems beyond our own.

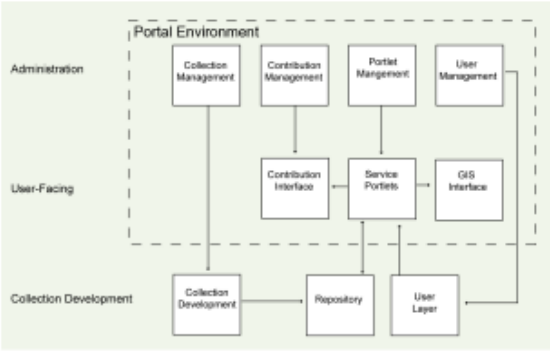

Figure 1: Overall component architecture in our model.

2. References

[1] Krowne, A. An OAI transformation-based architecture for digital library collection management (whitepaper), 2006.
[2] Krowne, A. OXF: The OAI/OCKHAM transformations framework (whitepaper), 2006.
[3] Morgan, E. L., Frumkin, J., and Fox, E. A. The OCKHAM initiative - building component-based digital library services and
collections. D-Lib Magazine 10, 11 (November 2004).